Emergent Roles in Computer-Mediated Synchronous Virtual Groups

Jordan B. Barlow
University of St. Thomas
jordan.barlow@stthomas.edu

Abstract

Understanding how the emergence of roles affects virtual group outcomes is important for organizations that increasingly use virtual work for decision-making and other tasks. Using role theory and speech act theory, this paper describes two studies conducted to understand the emergence of communication roles and their impacts on virtual group dynamics. Study 1 explores the emergence of roles in computer-mediated decision-making groups, using chat transcripts from a lab experiment. Study 2 further explores and validates the emergence of these roles, using a text mining technique to automate speech act analysis, and tests how these roles affect group perceptions of trust, communication, and performance. The paper concludes with a discussion of implications for practice and future research on the effects of emerging roles and their interactions.

1. Introduction

Roles are created or emerge in every situation where humans interact with each other [5]. Roles can be formal or informal; they can form naturally or by appointment. People can fill roles without formal recognition or even recognizing the roles themselves [14]. Roles often emerge as participants interact with each other and develop norms of behavior [33]. These roles emerge not only as a result of differing individual characteristics [e.g., 35], but also as a result of social interaction [33, 11].

Role emergence is particularly interesting in settings where no formal roles are defined, such as self-managing virtual groups. While past research on face-to-face decision groups gives some insight into the issue of emergent roles [e.g., 5], virtual decision-making groups have some distinct characteristics [1].

Group members have a need for role definition, particularly in virtual groups where roles tend to be more ambiguous [12]. Role clarity leads to improved group identity [12]; improved clarity and understanding of roles could also lead to other improved group outcomes. For example, recent research shows that the emergence of certain roles in online study groups can ultimately affect the grades of the group members [7]. Given that roles drive individual outcomes in study groups, emerging roles in virtual decision-making groups could similarly affect group decisions and other outcomes such as performance, communication effectiveness, or trust.

Typically, roles are defined as norms or patterns of behavior [5]. In many computer-mediated environments, behavior takes place through written language [16, 19]. Thus, roles in virtual settings are often based on communication patterns. Such patterns could include the types of speech acts used, the level of participation in group communication, and which participants communicate with others. While earlier research on role emergence in computer-mediated communication (CMC) has examined roles arising from behavioral expectation in online communities and other virtual asynchronous settings [13, 34, 3, 24, 11], fewer studies have focused on role emergence in synchronous online communication and virtual groups decision making.

Further, these studies have been descriptive and have not examined the outcomes (nor antecedents) of emergent roles in these groups. Research is only just starting to show that the configuration of roles affects group outcomes [7, 24]. Such research has so far focused on the roles of group members working asynchronously, with no research on the group outcomes resulting from emerging roles in synchronous, interactive virtual groups.

The current research, grounded in role theory, speech act theory, and theories of group communication, trust, and performance, uses computer-mediated discourse analysis (CMDA) tools and data mining techniques (cluster analysis and text mining) to address the following research questions:

RQ1. What roles emerge in computer-mediated decision-making tasks when groups use synchronous communication?

RQ2. What group outcomes are significantly related to emergent roles in synchronous, computer-mediated decision-making tasks?

This study provides three distinct contributions.

URI: https://hdl.handle.net/10125/59516
ISBN: 978-0-9981331-2-6

First, this paper is the first to examine roles that emerge based on communication patterns in synchronous contexts, specifically in the context of group decision making. Second, this research is one of few empirical studies to use a theory-based approach in analyzing communication examining emergent roles. Of these theory-based studies, this research is the only study to examine the effects of role emergence on group outcomes. Third, this research demonstrates the utility of various predictive analytic techniques to examine virtual group roles and outcomes; this study is one of the first to automate computer-mediated discourse analysis methods using text mining.

2. Study 1

2.1. Theoretical Background

Role theory defines roles simply as characteristic behavior patterns [5]. Role theory assumes that people hold social positions while interacting, and that they hold expectations for their own and others' behavior [5]. Further, the concept of roles provides a means for explaining the relationship between individuals and social structures by explaining how social structures are built upon individual patterns of action [6]. In online groups and communities, where formal roles are often not defined, an interactionist or emergent view of roles is especially appropriate because individuals are likely to respond to the needs of the group or community and adopt the role that seems most appropriate at the time [11]. In many computer-mediated environments, behavior takes place through written language [16, 19]; in accordance with role theory, such written language is the basis for roles that emerge in such environments.

Several researchers have examined the roles that people assume as they communicate through computer networks [13, 34, 3, 24]. These works have identified and analyzed different types of roles that emerge in asynchronous communication, including Usenet [34] and Wikipedia [14, 3, 24]. For example, a myriad of roles emerge in asynchronous collaborative learning environments, such as 'encourager,' 'dominator,' and 'fellow-traveler' [7]. Researchers have used social network analysis [13] and visualization methods [34] to understand emerging roles in Usenet groups such as 'discussion people' and 'answer people.' These roles are specific to asynchronous communication and do not fit well for synchronous decision tasks. Synchronous communication requires participants to be actively involved, rather than selectively deciding when to participate and with whom to interact.

In addition to role theory, this study draws from speech act theory [4], which posits that speech uttered in a particular way constitutes an act [19, 32]. Speech acts do not simply describe the world—they "bring about change to the world" [2: p. 252]. Speech act theory is particularly suited to virtual communication because in many computer-mediated environments, behavior is performed primarily (and sometimes only) through language [16, 19]. In other words, to understand the behavior of individuals in mediated groups, one must understand each speech act as constituting a certain type of behavior.

Surprisingly, no prior studies on role emergence in CMC [e.g., 13, 34, 3, 24], could be found that are grounded in speech act theory. Hence, they do not focus heavily on message content to understand roles, but rather focus on how much and with whom individuals communicate. However, according to speech act theory, behavior can only be truly understood by understanding the meaning of communicative acts. Using these perspectives from role theory and speech act theory, this research proposes that roles in virtual group decision-making tasks can be detected by studying the speech acts of individuals participating in the task.

2.2. Methodology and analysis

This study uses computer-mediated discourse analysis (CMDA) to analyze 26 group chat transcripts taken from a lab experiment. Participants in the lab were 130 undergraduate students (62% male) from various business courses, working in 26 groups of five. Participants used Gmail Chat to communicate with group members and complete a college admissions task as groups. Groups were instructed to make admissions decisions regarding a set of five possible candidates. To encourage discussion, groups were not allowed to admit more than three of the five candidates and were given both common and unique information, along with a 20-minute time constraint. This task has been used in previous research on virtual decision-making [30].

Chat transcripts of each session were downloaded from the Gmail accounts used in the study. The 26 transcripts contained 2996 messages, with an average of 115.23 messages per transcript.

Participation analysis is a common CMDA method used to detect roles in asynchronous CMC [e.g., 13, 34]. In this research, the number of messages by participant, the percentage of the group's messages by each participant, and the length (in words) of each message are analyzed.

Previous research on roles in CMC generally

combines participation analysis with another type of analysis—often, social network analysis because of the asynchronous nature of the studies [e.g., 14]. In this synchronous context, speech act analysis is a better fit with the theoretical reasoning and methodological design of the research.

In speech act analysis, a technique based on speech act theory [4, 32], researchers read every utterance in a communication transcript and assign it to a speech act category describing the behavior portrayed by the utterance (e.g., "claim", "question", "reaction"). To apply speech act analysis, taxonomies have been created by linguists. However, due to challenges of previous speech act taxonomies, Herring, et al. [17] developed a consolidated and simplified taxonomy specifically for the analysis of CMC. This CMC act taxonomy, which consists of 16 types of speech acts, was tested by Herring and her colleagues on blog posts, threaded bulletin boards, and synchronous chat. They found it to be easily applied and interpreted. The CMC act taxonomy has been used in previous CMC research [20].

Each of the 2996 messages of the sample were classified according to its relevant speech act type using the CMC act taxonomy. Some messages contained multiple speech acts; these were divided and coded as separate messages, resulting in a classification of 3055 total speech acts. Because speech act analysis can be subjective, a second coder participated in the analysis to establish inter-rater reliability. The two coders jointly coded one transcript of 170 speech acts to ensure mutual understanding of the classification scheme. Next, they independently coded four more transcripts (458 speech acts, about 15 percent of the total sample), and obtained a Cohen's kappa reliability score of 0.708, indicating substantial agreement between raters [c.f. 22]. The rest of the messages were then coded solely by one author.

2.3. Results

Participation analysis shows that messages in this type of conversation are typically short (mean = 5.60 words per message), with one-third of messages being only one or two words in length. The average number of messages per participants was 23.05. Results from the speech act analysis (shown in Table 1) indicate that most of the acts performed by participants through their communication were sharing information, making claims about information, agreeing with others, and asking questions. These speech acts are consistent with the type of task completed by the group.

Based on the summary results of the participation and speech act analyses and a qualitative reading of the text, a set of six relevant factors of role behavior was chosen to be used in a cluster analysis: two participation factors and four speech act factors.

Table 1. Speech act summary statistics

Category	%	Category	%
INFORM	22.8%	MANAGE	4.2%
CLAIM	19.9%	REQUEST	3.4%
ACCEPT	16.2%	REPAIR	2.6%
INQUIRE	9.9%	ELABORATE	2.5%
REACT	7.7%	GREET	1.7%
INVITE	7.0%	Others (< 1%)	2.1%

Though three participation analysis metrics were collected, only two were used as factors in the cluster analysis: percentage of group messages contributed by an individual, and average message length. The number of total messages from an individual was excluded for two reasons. First, this metric is redundant with the percentage of group messages contributed by an individual. Second, the percentage of group messages was preferable above total messages because it accounts for between-group differences in the number of messages typed.

The four speech act factors used were the following: information shared (percentage of an individual's speech acts classified as 'inform' speech acts), opinions shared (percentage of speech acts classified as 'claim' speech acts), agreement with others (percentage of speech acts classified as 'accept' speech acts), and amount of discussion guiding (percentage of speech acts classified as 'inquire,' 'manage, 'request,' or 'direct' speech acts). Each of these four factors accounted for over 10 percent of the overall number of speech acts, respectively. The remaining speech act types were not used frequently enough to make meaningful comparisons and were not as relevant to decision-making tasks. The 'inquire,' 'manage,' 'request,' and 'direct' speech acts were combined because these were strikingly similar upon reviewing them in the transcripts. Each of these types of speech acts is related to directing the communication of the group.

Using these six factors, cluster analysis was used for role detection, as done in previous CMC role research [24]. Ward's method of hierarchical clustering was first used to identify the proper number of clusters. The results suggested four distinct groups of participants in the sample based on the six factors. Then, non-hierarchical (k-means) clustering was used to refine the original clusters. The means of each of the six factors for each of the

four clusters are shown in Table 2. Distinguishing values for each cluster are highlighted in light gray (low values) and dark gray (high values).

Table 2. Study 1 cluster analysis summary

	% msg	Len	Inform	Claim	Accept	Guide
S	15.4%	5.1	45.8%	9.6%	20.1%	9.4%
L	16.9%	6.5	24.5%	18.0%	29.8%	10.8%
M	22.4%	5.6	20.6%	15.9%	14.7%	26.5%
O	24.8%	6.1	13.9%	33.0%	8.7%	14.9%

S = Sharers; L = Listeners; M = Managers; O = Opinionaters

% msg = % of group's messages communicated by the individual

Len = Average message length for the individual

Inform / Claim / Accept = % of an individual's messages classified as a that speech act

Guide = % of an individual's messages classified as 'inquire,' 'manage,' 'request,' or 'direct' speech act

The largest cluster (44 individuals) was the cluster of *managers*, those who directed the conversation of the group by asking questions and requesting information. Managers tended to have more than the average number of messages, with moderate amounts of sharing, claiming, and agreeing. Often, groups had multiple managers as individuals worked together to move the work forward.

The next largest group (37 individuals) is the *listeners*, those who participated less in discussion and primarily agreed with others. While the listeners tended to share some information and make claims when they did speak, these individuals did not guide the discussion.

Fewer individuals (24) filled the role of *sharers*, those who mostly shared facts about the candidates but did not interject many opinions. They answered questions and shared what they knew so that the group would have enough information to make a good decision. However, they had fewer, shorter messages and usually did not propose solutions.

The last emerging role (25 individuals) was that of *opinionaters*, those who primarily tried to convince others of their own opinions. Most of the individuals in this role shared a variety of opinions, but were slow to accept the opinions of others, some even showing high tendencies toward stubbornness in the transcripts. They used many messages, often long, to persuade their teammates of their opinions.

3. Study 2

3.1. Theory and hypothesis development

This study further examines the four roles found in Study 1 and compares these roles against various outcomes. Previous research indicates that inclusion or exclusion of certain types of individuals can affect group processes and outcomes [15]. However, until recently, little research was done in exploring relationships between roles and outcomes [33]. Scholars have called for more research examining how the presence or absence of roles affects individual and group outcomes [14].

Initial research on role outcomes has shown some interesting results. The presence of certain roles in collaborative learning environments affects group project grades [7]; the quality of a Wikipedia article depends on the roles of the contributors and how they collaborate with each other [24]. Such relationships between roles and outcomes should be examined for virtual decision-making groups in synchronous tasks. Further, while these studies have examined group performance, it is important to understand other aspects of group dynamics, such as member support and group interaction and well-being [26].

Several outcomes could be affected by the emergence of roles in computer-mediated decision-making groups, but this study focuses on group member perceptions of trust, communication effectiveness, and individual performance. Role theory states that as part of the role development process, people develop expectations and opinions about the behavior of themselves and others. Thus, the roles that emerge in computer-mediated tasks should be accompanied by distinct perceptions and expectations of group members toward the specific roles. That is, for an individual in a synchronous computer-mediated group, how does the role assumed by the individual affect other group member perceptions toward the individual?

3.1.1. Trust. First, trust is based on the assumption that others will behave in the way that they are expected to [21]. Because roles often emerge as a result of the interaction of individuals and the resulting expectation of which members perform which roles, trust is integrally related to the formation and display of role behaviors.

Research on virtual work suggests that trust is built on communication patterns between people [31]. In distributed groups, there is little observation of others' behavior on which to base opinions of trust, so group members must develop trust opinions by observing the way people communicate. Because group members with different roles exhibit different behaviors through communication patterns, it follows that the perceptions of trust toward these roles may be different. Further, because there are fewer social cues to interpret in virtual settings, group members must look to behavioral (i.e., communication)

patterns in order to develop a sense of how much they trust their fellow group members [31]. Research on credibility of online communication indicates that in online settings, people gauge the credibility (an important component of trust) of an argument based on whether they perceive the communication to be fact-based or opinion-based [27].

Hypothesis 1. Role emergence is related to trust, such that one or more of the emergent roles will be perceived as significantly more trustworthy than other roles.

3.1.2. Communication effectiveness. Similarly, because the various roles in a group have different communication patterns, perceptions of the communicative effectiveness of the patterns may vary among the group. That is, group members may perceive one pattern of communication as a more effective way to communicate than other patterns of communication. Research on communication in computer-mediated decision-making indicates that communication is seen as effective when it is straightforward; when excessive time is needed to make a decision, communication is seen as less effective [8].

Hypothesis 2. Role emergence is related to communication effectiveness, such that one or more of the emergent roles will be perceived as communicating more effectively than other roles.

3.1.3. Perceived performance. Finally, the patterns of communication exhibited by the various roles will likely lead to differences in the perceived performance of the individual holding the role. Research shows that virtual group members base perceptions of others' performance on perceptions of trust and communication skills [31]. Thus, as the different levels of trust and perceived communication effectiveness differ by role, group members will perceive the individuals in different roles as performing to a different degree. Further, research from psychology indicates that individuals who are quieter in conversation are often perceived to be less intelligent or inferior performers than others who communicate more freely [29].

Hypothesis 3. Role emergence is related to perceived performance, such that one or more of the emergent roles will be perceived as performing at a higher level than other roles.

3.2. Methodology and analysis

Chat transcripts and survey responses from 86 student groups performing tasks during lab sessions in a large university business school were used to test the hypotheses. Participants in the labs were 324 students from various business courses (57% male), working in randomly assigned groups of three to five.

Group members communicated using only Gmail Chat. Groups completed the university admissions task, similar to the task performed by group members in Study 1. In this instance, participants were instructed to make university admissions decisions regarding a set of four possible candidates; groups could admit up to two candidates in a 20-minute period.

At the end of the task, participants completed an online survey asking for demographic information and perceptions of fellow group members. Group perceptions of individual trustworthiness, communication, and performance were collected following the procedures of Sarker, et al. [31], where each group member rated every other member of the group on each construct on a Likert scale. Because trust and communication are inherently relational, rather than inherent individual characteristics, a social-network-based measure is more appropriate than traditional attribute-based measures [31].

CMDA methods were again used in coding and analyzing 6206 messages from the Gmail Chat transcripts. To determine the roles of the study participants, the same two CMDA methods as in Study 1—participation analysis and speech act analysis—were used. Following the participation analysis procedures of Study 1, the number of messages by participant, the percentage of the group's messages by each participant, and the length (in words) of each message were collected.

Previous studies employing speech act analysis [20] have required the manual coding of hundreds of messages, as done in Study 1. Because coding messages manually is time consuming, the possibility of automating speech act analysis through text mining is appealing both for the current study as well as for future research. Thus, in addition to examining emergent roles and their effects on the dynamics of virtual group work, this study also contributes by showing how text mining techniques can be used to automate computer-mediated discourse analysis in future research studies.

To build a model that would automatically assign speech act types to messages, the LightSIDE machine learning tool [25] was used. This tool transforms text into a set of features that can be analyzed with data mining algorithms. The dataset of speech acts from Study 1 was used as a training set in the model, given that Study 2 used a similar task and context. After experimenting with several models in the LightSIDE tool, the following options produced a feature set with highest accuracy: (1) Stopwords were retained

in the dataset. In all cases, removing stopwords reduced the predictive accuracy of the model. (2) Stemming was not used, as stemming also did not improve predictive accuracy in the model. (3) Punctuation was included in the dataset, as this tended to improve the accuracy. It may be that in the case of determining speech act types, it is important to consider variation in word forms, common words, and punctuation. (4) Spell-check tools were run on the training text using Microsoft Excel, which slightly improved the accuracy of the model. To be consistent, spell-check was also used on the data for the current study. (5) The logistic regression function was used to classify the speech acts in the training set, resulting in higher accuracy than SVM, Naïve Bayes, and other models. (6) 10-fold cross-validation was used to determine accuracy and kappa values.

Using these specifications, the model achieved (an unfortunately low) predictive accuracy of 65.7 percent. One of the reasons speech act analysis has not been automated in the past is due to low predictive power of mining algorithms to capture the complex and subjective nature of language and assign one of 16 values that are often very similar. However, higher accuracy and confidence can be achieved, making automation feasible and desirable, when only a subset of the speech act types need to be predicted.

Because the dataset for Study 2 was similar to Study 1 (percentages for speech act categories were similar to those shown in Table 1), it was known that the most common and relevant speech act types to be studied were 'inform,' 'claim,' 'accept,' and a combination of 'inquire,' 'manage,' 'request,' and 'direct'. The remaining speech act types were not used frequently in the training set, nor expected to be used in the current dataset. Thus, including these speech act types would only serve to make machine learning more difficult.

Thus, the training set was adapted so that all 'inquire,' 'manage,' 'request,' and 'direct' speech acts were classified simply as 'manage.' All other speech act types besides 'inform,' 'claim,' and 'accept' were classified in the training set as 'other.' With the need to classify into one of five classes rather than sixteen, the logistic regression algorithm was able to classify the training data with 74 percent accuracy. Importantly, the model resulted in a kappa reliability score of 0.674, similar to the value of 0.708 obtained in Study 1 using human raters. Because speech act analysis is subjective even with human raters, and a kappa value indicating substantial agreement was obtained [22], the text mining model appeared sufficient to classify the data.

A cluster analysis of participants was used to determine roles, as done in previous CMC role research [24] and in Study 1. The same set of features were used. The EM algorithm was used to cluster the data into four clusters. Different numbers of clusters (from 2 to 7) were attempted to ensure that another clustering solution would not be a better fit for the data. SSE and log likelihood values increased nearly linearly across the various solutions, meaning that no single solution, statistically speaking, was superior. However, the four-cluster solution was highly interpretable and similar in results and interpretation to the solution used in Study 1.

3.3. Results

The means of each of the features for each of the clusters are shown in Table 3. Distinguishing values are highlighted in light (low) and dark (high) gray.

Table 3. Study 2 cluster analysis summary

	% msg	Len	Inform	Claim	Accept	Guide
S	30.7%	7.4	28.4%	28.3%	9.2%	20.1%
L	19.7%	5.1	29.2%	18.3%	22.4%	6.7%
M	28.2%	6.9	13.4%	26.6%	12.9%	34.3%
O	24.5%	9.1	14.4%	48.8%	13.1%	11.8%

S = Sharers; L = Listeners; M = Managers; O = Opinionaaters

% msg = % of group's messages communicated by the individual

Len = Average message length for the individual

Inform / Claim / Accept = % of an individual's messages classified as a that speech act

Guide = % of an individual's messages classified as 'inquire,' 'manage,' 'request,' or 'direct' speech act

These clusters have some general similarity to the clusters found in Study 1. The largest cluster (133 individuals) was the cluster of *sharers*, those who mostly shared facts about the candidates. The individuals filling this role generally contributed moderately and were not extreme in the number or length of messages or in opinions shared. This role was somewhat different from the corresponding role in Study 1, where sharers had fewer, shorter messages and made fewer than average claims.

The next largest group (75 individuals) is the *listeners*, those who participated less in discussion and primarily agreed with others. In both studies, listeners tend to share information when they speak, but do not make efforts to guide the discussion of the group. Listeners in Study 2 contributed even fewer opinions and had shorter messages.

Fewer individuals (53) filled the role of *managers*, those who directed the conversation by asking questions and requesting information. In both studies, the managers tend to have moderate amounts of claiming and agreeing. In Study 2, managers

shared less information and were more moderate in the amounts of messages contributed.

The last emerging role (63 individuals) was that of *opinionaters*, those who primarily tried to convince others of their own opinions. Many individuals in this role shared a high number of opinions but were slow to accept the opinions of others. Again, these individuals used long messages to try to persuade their teammates of their opinions but did not facilitate the discussion.

The final step was to find which variables have statistically significant relationships to the roles. To test the hypotheses, a separate analysis of variance (ANOVA) was run for each variable to determine statistically significant mean differences between the roles. Post hoc tests are used to determine which roles are significantly different on each variable. Specifically, the Tukey test was used because it corrects to avoid inflated Type I errors.

Perceptions of trust among group members were not significantly related to the emergent role of an individual in the task ($F = 1.144$, $p = 0.332$), providing no support for H1. Perceptions of communication effectiveness were found to be significantly different by role ($F = 3.571$, $p = 0.014$), lending support to H2. Perceptions of individual performance were also found to be significantly different by role ($F = 4.071$, $p = 0.007$), lending support to H3.

Post hoc tests revealed that sharers were perceived as communicating more effectively (M = 5.85) than listeners (M = 5.38; $p = 0.008$) by group members. No other pairwise comparisons were statistically significant. Sharers were also perceived as performing higher in the task (M = 5.86) than listeners (M = 5.41; $p = 0.005$), but again, no other pairwise comparisons were statistically significant.

4. Discussion

4.1. Interpretation of findings

Each of the roles detected in this study reflects distinct patterns of behavior exhibited by participants during a decision-making process in a text-based communication environment.

The most common role in the Study 1 groups was the manager. The managers in a group were those who showed leadership behaviors as they directed the discussion, asked questions, and structured the communication. Often, groups would have more than one manager as individuals worked together to move the discussion forward with limited time. Some groups even had three or four managers. This complex pattern of leading the discussion and communication is consistent with prior CMC research suggesting that a simple pattern of single or complementary leadership does not always hold in online contexts [e.g., 7, 14] and emergent leadership patterns are different in virtual groups than in face-to-face groups [28]. In online contexts, even when task-related, the simple emergence of one or even two distinct leaders is not a given [28]; rather, individuals participating in CMC show more complex patterns of leading and following by assuming different types of roles. Fewer managers emerged in the Study 2 groups, perhaps because the groups in that study were smaller; it may be that in groups of three or four, participants are more likely to equally share information (as "sharers") while in larger groups of five participants, group members feel the need to make sure the group is well managed.

Online decision-making groups also tend to have many listeners, those who were mostly passive and agreeing with others. In both studies, between 20 and 30 percent of participants emerged as listeners. This finding is consistent with research showing that social loafing tends to emerge in groups larger than two to three individuals [9]. These listeners were generally doing more than just lurking; when asked for information or an opinion they would share it. The amount of participation of these individuals, along with the high amount of agreement, showed a mostly passive role. In Study 1, listener messages were longer than average; in Study 2, they were shorter. Thus, message length can't be concluded as a distinguishing factor of this type of role. In other words, listeners share few message, with little discussion guiding, regardless of the length of the actual messages.

In Study 1, sharers had more distinctive characteristics. In Study 2, with the larger sample size, the results showed a more intuitive and stable view of the sharers cluster. Across the two studies, the sharer role is one who participates more actively than listeners, but who does not take a strong role as a leader or opinionater in the discussion. This role can take many different forms, and the higher the sample size (as in Study 2), the more likely it is to see percentages of speech act types regress toward the mean. Study 2 showed particular characteristics for the other three roles but showed sharers as having an average amount of each speech act. In Study 2, the number of sharers was larger than any other role. In Study 1, it happened to be that sharers contributed fewer of the messages, but there were also more managers on average, per group. The role of sharer seems to be the most flexible of the group and can manifest in many ways, but with a larger study, the

number of different types of sharers average out in the summary statistics.

Opinionaters were those whose primary goal was to convince others of their own opinions. In both studies, about 20 percent of participants assumed this role. Most of the individuals in this role shared a variety of opinions but were slow to moderate in accepting the opinions of others. While the number of speech acts categorized as 'reject' in Study 1 was low overall, speech acts connoting a rejection of others' ideas was highest in this cluster. Further, opinionaters shared relatively low amounts of factual information with the group, relying on opinion more than evidence to argue a case.

A variety of all four roles emerged in groups at all levels of decision quality; that is, groups who performed better on the task did not have a unique role structure. The emergence of roles is affected by individual characteristics, but the transcripts indicated that in some cases, people assumed roles as a result of the group interaction. For example, if a group happened to have many managers or opinionaters dominating the discussion, others who may have assumed these roles became more passive and assumed a listener or sharer role instead.

The results of this research demonstrate that emergent roles have significant effects on perceptions of group members toward individuals. Sharers were perceived as high performers and effective communicators in Study 2. Group members perceived sharer, those who were most active without assuming strong opinions or leadership behaviors, as being the best communicators and performers.

On the contrary, listeners are those individuals who are seen as less effective communicators and performers than sharers. These individuals are those who contribute less to the discussion, with a large share of their communication being only to agree with or accept the information and opinions put forth by others. It may be that listeners are sometimes seen as social loafers by their fellow group members [9]. However, it should be noted that while listeners are seen as significantly less effective at communicating and performing with the group than sharers, the average ratings given to listeners are still positive (i.e., greater than 5 on a seven-point Likert scale). In other words, listeners can still be seen as somewhat effective communicators and performers, but are not perceived as highly as sharers.

Neither managers nor opinionaters were seen as significantly more (nor less) trustworthy, effective in communication, or likely to perform well. Perceptions of these roles may depend on whether participants agreed with managers' style of guidance or the opinions shared by opinionaters.

4.2. Practical implications

Awareness of the unique task-based roles that emerge during virtual decision-making tasks, and their effects on group dynamics, is important for groups that commonly use technology. Understanding the types of behaviors that emerge in these settings and how group members will perceive these patterns of behavior is useful to better structure and facilitate communication among individuals. For example, organizations may try to find people to assume the role of sharer in groups and discourage the assumption of the listener role to have a positive effect on group work. Future research should continue to examine the outcomes of the interaction of such roles to help organizations understand the best way to structure groups in terms of roles.

Understanding the emergent roles of virtual decision-making groups can also inform the design of information technology to support such virtual groups. Systems can be created that facilitate communication and coordination based on communication patterns of individuals within the group [10], and such interventions may lead to improved group outcomes.

4.3. Theoretical contribution

This research also has implications for research on the relationship of roles to group outcomes. Previous research indicates that inclusion or exclusion of certain types of individuals can affect group processes and outcomes [15]. However, until recently, little research was done in exploring relationships between roles and group-level outcomes [33], and scholars have called for more research examining how the roles affect group outcomes [14]. This research was the first step in understanding the relationship of roles in computer-mediated synchronous tasks to groupwork outcomes at the individual level. Future research will examine the effects at the group level.

Several other outcomes could be affected by the emergence of roles. For example, roles may affect the extent to which individuals change their decisions to conform to the group. Future research should continue to examine additional group outcomes influenced by these roles. Further, if roles affect group outcomes, and group performance is correlated with individual characteristics such as gender, educational background, and culture [e.g., 18, 35], behavioral roles may act as a mediator between individual characteristics and group outcomes.

Traditional role research has viewed roles as

somewhat static within individuals. Indeed, roles are known to be influenced by individual characteristics [33]. However, research also shows that individual roles are adapted in real-time depending on the context and situational demands [33]. This opens up the possibility that individuals and organizations can use interventions to change the emerging roles in group contexts.

Researchers have developed interventions to improve group collaboration, and more research is needed to understand how these could affect role emergence. For example, technological interventions have been designed to give real-time feedback to group members about the type of language they use during group discussion [23]. These intervention systems have the power to affect roles that emerge during computer-mediated tasks.

Finally, this study has implications for the methods in further research on roles in virtual settings. This research successfully used speech act analysis, a form of CMDA, to detect and interpret roles in synchronous computer-mediated communication. Study 2 used text mining to partially automate speech act analysis as part of the role detection process. Further research should build on this work by continuing to use text mining, speech act analysis, and other CMDA methods that are theoretically related to the research question regarding roles in online communication. A study of speech acts directly reflects participant behavior in computer-mediated settings, and such techniques should be applied in studies of online behavior.

4.4. Limitations

The results of this research should be considered in light of some limitations. First, the transcripts for this study were taken from a lab experiment using student participants. While the results are useful for a variety of settings, including student groups, careful consideration should be used in applying the results to other contexts where individuals have different characteristics than students. Because roles and behavior are dependent on a multitude of factors, including gender, education, and culture [18], roles may differ in other contexts outside of student groups. However, because the results demonstrate general human behavior, similar results could potentially remain valid in other groups. Further, in this study participants were anonymous; in other settings, interactions prior to working on a task could affect group perceptions.

Care should also be taken in generalizing the results to other types of CMC or other types of group tasks—this study only examined text-based chat for decision-making groups. Further, the results regarding sharers as effective communicators and performers were found in Study 2, where sharers were more well-rounded than in Study 1. However, a main contribution of this research is a demonstration of an appropriate technique for detecting roles in CMC. This technique can be applied to CMC in other contexts in future research.

5. References

[1] A. Abbasi and H. Chen, "CyberGate: A design framework and system for text analysis of computer-mediated communication", MIS Quarterly, 32 (2008), pp. 811-837.

[2] P. J. Ågerfalk, "Getting pragmatic", European Journal of Information Systems, 19 (2010), pp. 251-256.

[3] D. Anthony, S. W. Smith and T. Williamson, "Reputation and reliability in collective goods: The case of the online encyclopedia Wikipedia", Rationality and Society, 21 (2009), pp. 283-306.

[4] J. L. Austin, *How to Do Things with Words*, Harvard University Press, Cambridge, MA, USA, 1962.

[5] B. J. Biddle, "Recent developments in role theory", Annual Review of Sociology, 12 (1986), pp. 67-92.

[6] P. L. Callero, "From role-playing to role-using: Understanding role as resource", Social Psychology Quarterly, 57 (1994), pp. 228-243.

[7] C.-K. Chang, G.-D. Chen and C.-Y. Wang, "Statistical model for predicting roles and effects in learning community", Behaviour & Information Technology, 30 (2011), pp. 101-111.

[8] L. Chidambaram and B. Jones, "Impact of communication medium and computer support on group perceptions and performance: A comparison of face-to-face and dispersed meetings", MIS Quarterly, 17 (1993), pp. 465-491.

[9] L. Chidambaram and L. L. Tung, "Is out of sight, out of mind? An empirical study of social loafing in technology-supported groups", Information Systems Research, 16 (2005), pp. 149-168.

[10] T. Erickson and W. A. Kellogg, "Social translucence: An approach to designing systems that support social processes", ACM Transactions on Computer-Human Interaction, 7 (2000), pp. 59-83.

[11] S. Faraj, S. L. Jarvenpaa and A. Majchrzak, "Knowledge collaboration in online communities", Organization Science, 22 (2011), pp. 1224-1239.

[12] C. M. Fiol and E. J. O'Connor, "Identification in face-to-face, hybrid, and pure virtual teams: Untangling the contradictions", Organization Science, 16 (2005), pp. 19-32.

[13] D. Fisher, M. A. Smith and H. T. Welser, *You are who you talk to: Detecting roles in Usenet newsgroups*, *Proc. 39th HICSS*, Computer Society Press, January 4-7, 2006, 2006, pp. 1-10.

[14] E. Gleave, H. T. Welser, T. M. Lento and M. A. Smith, *A conceptual and operational definition of 'social role' in online community*, *Proc. 43rd HICSS*, Computer Society Press, January 5-8, 2009, 2009, pp. 1-11.

[15] R. A. Guzzo and M. W. Dickson, "Teams in organizations: Recent research on performance and effectiveness", Annual Review of Psychology, 47 (1996), pp. 307-338.

[16] S. C. Herring, *Computer-mediated discourse analysis: An approach to researching online behavior*, in S. A. Barab, R. Kling and J. H. Gray, eds., *Designing for virtual communities in the service of learning*, Cambridge University Press, New York, NY, USA, 2004, pp. 338-376.

[17] S. C. Herring, A. Das and S. Penumarthy, *CMC act taxonomy*, http://ella.slis.indiana.edu/~herring/cmc.acts.html, 2005.

[18] P. Hinds, L. Liu and J. Lyon, "Putting the global in global work: An intercultural lens on the practice of cross-national collaboration", The Academy of Management Annals, 5 (2011), pp. 135-188.

[19] M. A. Janson and C. C. Woo, "A speech act lexicon: An alternative use of speech act theory in information systems", Information Systems Journal, 6 (1996), pp. 301-329.

[20] S. Kapidzic and S. C. Herring, "Gender, communication, and self-presentation in teen chatrooms revisited: Have patterns changed?", Journal of Computer-Mediated Communication, 17 (2011), pp. 39-59.

[21] K. Knoll and S. L. Jarvenpaa, *Learning virtual team collaboration*, in R. H. Sprague, ed., *28th Annual Hawaii International Conference on System Sciences*, IEEE Computer Society Press, 1995, pp. 92-101.

[22] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data", Biometrics, 33 (1977), pp. 159-174.

[23] G. Leshed, D. Perez, J. T. Hancock, D. Cosley, J. Birnholtz, S. Lee, P. L. McLeod and G. Gay, *Visualizing real-time language-based feedback on teamwork behavior in computer-mediated groups*, *CHI 2009*, ACM, 2009, pp. 537-546.

[24] J. Liu and S. Ram, "Who does what: Collaboration patterns in the Wikipedia and their impact on article quality", ACM Transactions on Management Information Systems, 2 (2011), pp. article 11.

[25] E. Mayfield and C. P. Rosé, *LightSIDE: Open source machine learning for text*, in M. D. Shermis and J. Burstein, eds., *Handbook on Automated Essay Evaluation: Current Applications and New Directions*, Routledge Academic, 2013.

[26] J. E. McGrath, "Time, interaction, and performance (TIP): A theory of groups", Small Group Research, 22 (1991), pp. 147-174.

[27] M. J. Metzger, "Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research", Journal of the American Society for Information Science and Technology, 58 (2007), pp. 2078-2091.

[28] N. I. Misiolek and R. Heckman, *Patterns of emergent leadership in virtual teams*, *Proc. 38th HICSS*, IEEE Press, 2005, pp. 1-10.

[29] D. L. Paulhus and K. L. Morgan, "Perceptions of intelligence in leaderless groups: The dynamic effects of shyness and acquaintance", Journal of Personality and Social Psychology, 72 (1997), pp. 581-591.

[30] L. Robert and A. R. Dennis, "The paradox of richness: A cognitive model of media choice", IEEE Transactions on Professional Communication, 48 (2005), pp. 10-21.

[31] S. Sarker, M. Ahuja, S. Sarker and S. Kirkeby, "The role of communication and trust in global virtual teams: A social network perspective", Journal of Management Information Systems, 28 (2011), pp. 273-309.

[32] J. R. Searle, *A taxonomy of illocutionary acts*, in J. R. Searle, ed., *Expression and Meaning: Studies in the Theory of Speech Acts*, Cambridge University Press, Cambridge, UK, 1979, pp. 1-29.

[33] G. L. Stewart, I. S. Fulmer and M. R. Barrick, "An exploration of member roles as a multilevel linking mechanism for individual traits and team outcomes", Personnel Psychology, 58 (2005), pp. 343-365.

[34] H. T. Welser, E. Gleave, D. Fisher and M. A. Smith, "Visualizing the signatures of social roles in online discussion groups", Journal of Social Structure, 8 (2007), pp. 1-32.

[35] R. Zolin, P. Hinds, R. Fruchter and R. E. Levitt, "Interpersonal trust in cross-functional, geographically distributed work: A longitudinal study", Information and Organization, 14 (2004), pp. 1-26.